

CRISPR-Cas9 Mediated Knockout of NtAn1 to Enhance the Lipid Accumulation in Tobacco Seed for Biodiesel Production

Yinshuai Tian

Hebei University of Engineering

Xinanbei Liu

Sichuan University

Tingting Li

Sichuan University

Huan Qin

Sichuan University

Xiao Li

Sichuan University

Kai Chen

Hebei University of Engineering

Fang Chen

Sichuan University

Ying Xu (✉ xuying@scu.edu.cn)

Sichuan University

Research

Keywords: Tobacco, Seed lipid, NtAn1, CRISPR-Cas9, Biodiesel

Posted Date: June 24th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-37669/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Tobacco seed lipid is a promising non-edible feedstock for biodiesel production. In order to meet the increasing demand, achieving high seed lipid content is one of the major goals in tobacco seed production. The *TT8* gene and its homologs negatively regulate seed lipid accumulation in *Arabidopsis* and *Brassica* species. We speculated that manipulating the homolog genes of *TT8* in tobacco could enhance the accumulation of seed lipid.

Results: In this present study, we found that the *TT8* homolog genes in tobacco, *NtAn1a* and *NtAn1b*, were highly expressed in developing seed. Targeted mutagenesis of *NtAn1* genes were created by the CRISPR-Cas9 based gene editing technology. Due to the defect of PAs biosynthesis, mutant seeds showed a phenotype of yellow seed coat. Seed lipid accumulation was enhanced by about 18% and 15% in two targeted mutant lines, respectively. Protein content was also significantly increased in mutant seeds. In addition, the seed yield related traits were not affected by the targeted mutagenesis of *NtAn1* genes. Thus, the overall lipid productivity of the *NtAn1* knockout mutants were dramatically enhanced.

Conclusion: Tobacco *NtAn1* genes regulate both PAs and lipid accumulation in the process of seed development. Targeted mutagenesis of *NtAn1* genes could generate a yellow-seeded tobacco variety with high lipid and protein content. Furthermore, the present results revealed that CRISPR-Cas9 system could be employed in tobacco seed *de novo* domestication for biodiesel feedstock production.

Background

Due to the increasing concerns about climate change resulting from excessive consumption of petroleum products, biodiesel has attracted more and more attentions in the recent years. Tobacco (*Nicotiana tabacum* L.) is an oilseed plant with a high seed lipid content ranging from 36 to 41% of the seed dry weight [1, 2]. Recently, tobacco seed lipid had been demonstrated to be a promising feedstock for biodiesel production [3–6]. Furthermore, due to the rich carbohydrates, wide availability, and low cost, tobacco stalk could also be used for biofuel production [7]. Recently, a high seed yield tobacco variety, Solaris, had been bred by the Sunchem Holding Company for seed lipid feedstock production [8]. Life cycle analysis showed that the impacts created by the production of Solaris tobacco biodiesel were similar to those from other biodiesel plants [9]. Sustainable provision of feedstock is the key to sustainable biofuels [10]. In order to meet the huge demand for the feedstock of biodiesel production, achieving high tobacco seed lipid content is one of the main goals in the future.

Sucrose produced by photosynthetic tissues serves as a major carbon source for the synthesis of both seed storage compounds and generation of other seed components, such as mucilage and proanthocyanidins (PAs, also called as condensed tannins) in seed coat. Seed coat development competes for sucrose with reserve component synthesis in embryo and endosperm. Recently, studies had demonstrated that the amount of PAs in the seed coat is negatively correlated with the amount of lipid content in the embryo in *Arabidopsis* and rapeseed [11, 12]. Thus, strategies of manipulating the PAs biosynthesis pathway in seed coat could be used to increase the seed lipid content.

PAs deposit in the inner integument of the seed coat, and the oxidation of PAs in the process of seed maturation results in the formation of brown pigments that confer color to the mature seed [13]. Previous studies had demonstrated that the biosynthesis of PAs were mainly regulated at the transcription level by transcription factors belonging to the basic helix-loop-helix (bHLH), R2R3-MYB, and WD-repeat protein families [13, 14]. In *Arabidopsis*, *TT2*, *TT8*, and *TTG1*, which encode R2R3-MYB, bHLH, and WD40 repeat proteins, respectively, form a ternary complex to activate the expression of PA-specific genes during seed coat development [14, 15].

Previous studies revealed that *TT8* gene played a key regulation role in seed PAs biosynthesis. In *Arabidopsis*, the *TT8* gene is required for the expression of *DFR* and *BAN* genes in siliques and young seedlings [16]. Due to the defect of PAs synthesis, *Arabidopsis tt8* mutants showed the transparent testa phenotype [16]. In addition to *Arabidopsis*, natural mutation of *TT8* genes resulted in yellow seed coat trait in allotetraploid *Brassica juncea* [17]. Most recently, targeted mutation of the *TT8* homologs through CRISPR-Cas9 system in *Brassica napus* also generated yellow-seeded phenotype [12]. Homologs of *TT8* gene were also reported to be involved in PAs biosynthesis in diverse plant species, including *Medicago truncatula*, *Lotus corniculatus*, and *Raphanus sativus* [18–20].

In addition to its critical role in PAs biosynthesis regulation, the *Arabidopsis* TT8 protein could repress the seed lipid accumulation through inhibiting the expression of transcription factors, including *LEC1*, *LEC2*, and *FUS3*, which play the key roles in embryo development and seed lipid biosynthesis [21]. Furthermore, TT8 protein could directly repress the expression of genes encoding enzymes involved in fatty acid biosynthesis by binding to the promoter region [21]. Thus, mutation of *TT8* gene generated seed with a thinner seed coat, a reduced PAs content, and an increased content of lipid in *Arabidopsis* and *Brassica* species [12, 21, 22]. We suggested that *TT8* gene could be a promising target aimed at enhancing the lipid content for oilseed plant.

Due to the presence of PAs, tobacco seed shows a dark brown seed color. Thus we proposed that *TT8* homolog genes in tobacco might be an ideal candidate to create high lipid content seed for biodiesel production. Previous report had identified and characterized two *TT8* homolog genes in tobacco genome, *NtAn1a* and *NtAn1b*, originated from two ancestors of tobacco, *N. sylvestris* and *N. tomentosiformis*, respectively [23]. *NtAn1* genes were demonstrated to be involved in flower anthocyanin biosynthesis [23], however, if *NtAn1* genes regulate the accumulation of PAs and lipid during seed development has not been reported. In this present paper, the expression pattern of *NtAn1* genes during seed development were analyzed, and CRISPR-Cas9 system was applied to generate *NtAn1* knockout mutants. We found that targeted mutagenesis of *NtAn1* genes significantly enhanced the accumulation of seed lipid and protein. These results demonstrated that CRISPR-Cas9 mediated knockout of *NtAn1* genes is an efficient approach to improve lipid production in tobacco.

Results

NtAn1 genes were highly expressed in developing seed

Previous study reported that *Arabidopsis TT8* gene has two homologs in tobacco genome, *NtAn1a* and *NtAn1b* [23]. Tobacco is a natural allotetraploid plant, sequence analysis revealed that *NtAn1a* originated from *N. sylvestris*, whereas *NtAn1b* derived from *N. tomentosiformis*. qPCR analysis showed that *NtAn1a* and *NtAn1b* expressed at developing flowers with highest expression level in corolla limb, which were consistent with the function of flower flavonoid biosynthesis regulation [23]. However, we noticed that the transcript level of both *NtAn1a* and *NtAn1b* were relatively high in developing ovary, which indicated that *NtAn1* genes might play an important role in seed development.

To further confirm and characterize the expression pattern of *NtAn1a* and *NtAn1b*, their expression were assessed in different organs and various stages of seed development: 7, 14, 21, and 28 days after flowering (DAF). qRT-PCR results showed that *NtAn1a* and *NtAn1b* had a similar expression pattern, with the highest expression in developing seeds at 7 DAF and decreased at later stages (Fig. 1). The expression of both *NtAn1* genes exhibited a high expression level in flower, which were consistent with the previous results. However, low transcript levels were detected in root, leaf, and stem (Fig. 1). These results suggested that *NtAn1a* and *NtAn1b* might regulate PAs and lipid accumulation during seed development in a way like in *Arabidopsis* and rapeseed.

Targeted mutagenesis of *NtAn1* using the CRISPR-Cas9 system

NtAn1a and *NtAn1b* showed a high sequence identity of 92.95% and 90.36% at the nucleotide and protein levels, respectively (Additional file 1: Figure. S1 and Figure. S2). Furthermore, these two genes had similar expression patterns (Fig. 1). These results suggested that these two genes may have similar and redundant functions. Thus, two gRNAs recognizing both *NtAn1* genes were designed to effectively knockout of them, and both of the gRNAs targeting the second exon of the coding sequence (Fig. 2a). The CRISPR-Cas9 construct containing these two gRNAs, which driven by the *Arabidopsis* U6-26 and U6-29 promoter, respectively (Fig. 2b), was produced based on the CRISPR-Cas9 multiplex genome-editing vector [24]. The resulting construct was transformed into wild type (WT) tobacco plant using *Agrobacterium*-mediated leaf disc transformation method. Through kanamycin selection, 12 kanamycin resistance T0 transgenic plants were generated. The targeting region of both *NtAn1a* and *NtAn1b* were amplified by a pair of primers at the same time. Two homozygous mutant lines (*an1-1* and *an1-2*) were identified from the T0 transgenic plants by Sanger sequencing analysis of the gRNAs targeting region. The *an1-1* mutant had one base insertion at both gRNA targeting sites of the *NtAn1a* and *NtAn1b* genes, while the *an1-2* mutant line had a 105 base fragment deletion between the two gRNAs targeting sites (Fig. 2c and Fig. 2d). T-DNA free mutant plants were selected from the T1 progeny generated by the self-pollinated of the two independent T0 homozygous mutant lines. Twenty T-DNA free T1 generation plants from each mutant line were randomly selected for further analysis.

Mutation of *NtAn1* genes resulted in yellow seed coat

The formation of seed color in most plant species is due to the deposition of PAs within the endothelial layer of the inner integument of the seed coat [13]. Previous studies demonstrated that *TT8* played a key role in regulating PAs accumulation in various plants [18, 19]. In this paper, mutation of the *NtAn1*, the

TT8 homolog genes in tobacco, generated yellow-seeded phenotype (Fig. 3a). This indicated that targeted mutation of the *NtAn1* genes might disrupt the accumulation of PAs in tobacco seed coat. In order to check the PAs deposition visibly, the mutant tobacco seeds were dyed by DMACA reagent. The results further confirmed the defects of PAs accumulation in seed coat (Fig. 3b). The soluble and insoluble PAs contents were calculated quantitatively by spectrophotometric method. The results showed that the PAs were mainly stored in the insoluble form in tobacco seed coat, and targeted mutation of *NtAn1* genes led to the significant decreases in both soluble and insoluble PAs content compared with those in WT tobacco seed coat (Fig. 3c).

ANR (Anthocyanidin reductase) and LAR (Leucoanthocyanidin reductase) are two key enzymes participated in PAs biosynthesis. ANR converts anthocyanidins to the epicatechin [25], whereas LAR could reduce leucocyanidin to catechin [26]. Catechin and epicatechin are considered to be the main building blocks for PAs biosynthesis. The encoding genes of ANR and LAR were directly regulated by *TT8* gene in *Arabidopsis* [15]. The expression patterns of the homolog genes encoding these two enzymes during seed development were analyzed by qRT-PCR. Our results showed that both of these two genes had a similar expression profile as *NtAn1*, and the expression of them were significantly decreased in both *an1-1* and *an1-2* mutant seed (Fig. 3d). Taken together, these findings indicate that tobacco *NtAn1* regulated the accumulation of PAs in a similar way like in *Arabidopsis*, and mutation of the tobacco *NtAn1* genes could hinder the PAs deposition in the seed coat, which was consistent with the phenotypes observed in *tt8* mutant seed in *Arabidopsis* and other *Brassica* species [12, 17, 21, 22].

Targeted mutagenesis of *NtAn1* generated white flower

Previous study had demonstrated that the anthocyanin accumulation in transgenic tobacco flowers could be significantly elevated by overexpression of *NtAn1a* or *NtAn1b* gene [23]. The early biosynthesis genes (EBGs) and late biosynthesis genes (LBGs) in anthocyanin pathway, including *CHS*, *CHI*, *F3H*, *DFR*, and *ANS*, were dramatically induced by the overexpression of *NtAn1* genes [23]. In this present paper, we found that targeted mutagenesis of *NtAn1* genes generated white flower phenotype, which resulted from the defects in anthocyanin accumulation in the flower (Fig. 4a). *TT8* genes regulate the biosynthesis of anthocyanin by manipulating the expression of the LBGs in anthocyanin pathway. The expression level of the downstream genes at different flower development stages were analyzed by qRT-PCR. Our results revealed that the examined anthocyanin biosynthesis genes expressed at all three developmental stages with expression level peaking at the late stage in WT plant flower (Fig. 4b). The expression patterns of these genes were consistent well with those of *NtAn1* genes, which indicating the regulation relation between them [23]. By contrast, the expression of all examined anthocyanin biosynthesis genes at different developing stages were significantly repressed in *an1-1* mutant line (Fig. 4b). Taken together, our results demonstrated again that *NtAn1* genes played an essential role in the biosynthesis of anthocyanin in tobacco flower.

Targeted mutagenesis of *NtAn1* increased seed lipid and protein contents

Both natural and targeted mutation of *TT8* genes in *Arabidopsis* and *Brassica* species would result in a significant increases in seed lipid content. To characterize the effect of the *NtAn1* genes targeted mutation on tobacco seed lipid accumulation, the seed lipid content was analyzed by GC-MS method. The results showed that WT tobacco seed lipid content was about 38.77 μg per seed, and the lipid content was approximately 45.91 μg per seed in *an1-1* and 44.97 μg per seed in *an1-2* mutant line, increased significantly by 18.42% and 15.99% relative to the WT seeds, respectively (Fig. 5a). These results indicated that the *TT8* gene and its homologs regulated seed lipid accumulation in a conserved way among different plant species. Thus, the *TT8* homologs from other oilseed plant could be used as a target to enhance the seed lipid content.

In most oilseed crops, the seed lipid content is negatively correlated with the protein content. Surprisingly, The *BnTT8* mutant seeds showed simultaneous increases in both lipid and protein contents [12], which is different from the lower protein content in *Arabidopsis tt8* mutant seed [21]. These results indicated that the seed protein accumulation was regulated by different mechanism in *Arabidopsis* and *Brassica napus*. To determine the effect of *NtAn1* genes on tobacco seed protein accumulation, the seed protein contents in WT and *an1* mutant lines were examined using the Pierce BCA Protein Assay Kit. The results showed that WT tobacco seed protein content was about 32.79 μg per seed, and the protein content was elevated to 36.56 μg per seed in *an1-1* and to 35.97 μg per seed in *an1-2* mutant line, increased significantly by 11.50% and 9.70% relative to the WT seeds, respectively (Fig. 5b). These results were consistent with those in *Brassica napus*. The increased protein content of the *an1* mutant seeds could make the tobacco seed meal produced after lipid extraction to be more valuable for animal feed manufacture.

The property of biodiesel is partially determined by the fatty acid chain length, the position and number of the double bonds [27]. In *Arabidopsis* and *Brassica napus*, the mutation of *TT8* gene resulted in an alteration in the seed fatty acid profile, including increases in palmitic acid, linoleic acid, and linolenic acid, while decreases in stearic acid and oleic acid compared with the WT seeds [12, 21]. Fatty acid composition of tobacco seed lipid shows a main presence of palmitic acid, stearic acid, oleic acid, and linoleic acid [3]. Possibly due the differences in the fatty acid composition, targeted mutation of *NtAn1* genes just resulted in a significant decrease in stearic acid, and the other four main fatty acid components were not changed significantly compared with the WT tobacco seed (Table. 1). WT tobacco seed lipid shows a high ratio of linoleic acid (~ 72%), which would make biodiesel produced from tobacco seed lipid more susceptible to oxidation, and this will limit its use in traditional engine. High oleic acid content is a preferred trait for biodiesel feedstock production [27]. Our previous work had generated a high oleic acid tobacco variety through CRISPR-Cas9 mediated knockout of the *NtFAD2*-genes [28]. Next, a high lipid content tobacco seed with an ideal fatty acid profile could be expected by combining the high lipid content trait generated in this present study with the high oleic acid phenotype through hybridization or multiple gene editing.

Table 1
Fatty acid composition of the *an1* mutant lines

	Fatty acid composition (mol%)				
	16:0	18:0	18:1 ^{Δ9}	18:2 ^{Δ9,12}	18:3 ^{Δ9,12,15}
WT	11.14 ± 0.37	4.36 ± 0.36	12.86 ± 0.35	71.14 ± 1.29	0.70 ± 0.02
<i>an1-1</i>	11.37 ± 0.28	2.77 ± 0.24**	13.02 ± 0.38	72.69 ± 1.43	0.98 ± 0.03
<i>an1-2</i>	11.96 ± 0.39	2.35 ± 0.21**	12.77 ± 0.29	73.08 ± 1.39	0.73 ± 0.01

16:0, palmitic acid; 18:0, stearic acid; 18:1^{Δ9}, oleic acid; 18:2^{Δ9,12}, linoleic acid; 18:3^{Δ9,12,15}, linolenic acid. The data are the mean ± SD (n = 3).

Expression of genes involved in seed development and lipid biosynthesis were altered by *NtAn1* targeted mutation

In *Arabidopsis*, TT8 protein could repress the lipid biosynthesis pathway through directly binding to the promoter region of the critical transcriptional factors important for seed development, such as *LEAFY*, *COTYLEDON1* (*LEC1*), *LEC2*, and *FUSCA3* (*FUS3*). Thus, the *Arabidopsis tt8* mutant showed increased seed lipid content [21]. In *Brassica napus* L., The expression levels of several genes involved in fatty acid biosynthesis during seed development were increased in *BnTT8* mutant plants generated by CRISPR-Cas9 system [12]. In this present paper, targeted mutagenesis of *NtAn1* genes led to the increases of both lipid and protein content, we suggested that *NtAn1* genes might regulate seed development and storage component accumulation in a way similar to that in *Brassica napus*. To test this hypothesis, qRT-PCR was performed in the process of seed development to examine the expression of genes involved in seed development regulation and fatty acid biosynthesis, including *LEC1*, *LEC2*, *FUS3*, *KASI*, *PI-PKβ1*, and *BCCP2*. Our results showed that the expression of all examined genes were up-regulated in the *an1-1* mutant at one or two developmental stages (Fig. 6). For example, the *LEC1* and *FUS3* genes were significantly increased at 14 DAF and 21 DAF seeds, while the *KASI* and *PI-PKβ1* genes were up-regulated at 14 DAF stage (Fig. 6). The elevated expression levels of genes involved in seed development and lipid biosynthesis could explain the enhanced lipid content in mutant lines.

Seed yield related traits were not affected by the mutation of *NtAn1*

Besides the seed lipid content, seed yield is another important factor affecting the lipid yield. The seed yield related traits of the *an1-1* and *an1-2* mutant lines were also evaluated. The results showed that the yield related traits, including seed size (Fig. 7a and Fig. 7b), seed weight (Fig. 7c), fruit number per plant (Fig. 7d), and seed number per fruit (Fig. 7e) were not affected by targeted mutation of *NtAn1* genes. Thus, the targeted mutation of *NtAn1* genes could generate a useful tobacco variety with a high seed lipid yield and improved nutritional quality.

Discussion

TT8 homolog gene functions in a conserved way to regulated seed coat PAs and seed lipid accumulation

Because of the character of a thinner seed coat, a reduced PAs content, and an increased content of lipid, yellow-seeded phenotype is a preferred trait for numerous oil seed plant. Due to the accumulation of PAs in the seed coat, tobacco seed showed a black color. Previous studies had demonstrated that TT8 protein played a key regulation role in PAs biosynthesis in seed coat [29]. Two *TT8* homologs, *NtAn1a* and *NtAn1b*, had been identified in tobacco genome [23]. However, the functions of seed color formation and seed lipid accumulation of these two genes had not yet been characterized. In this paper, we found that *NtAn1* genes were highly expressed in developing seeds (Fig. 1), which indicated that they might play a role in seed development. A yellow seed phenotype was generated by CRISPR-Cas9 mediated targeted mutagenesis of tobacco *NtAn1* genes (Fig. 3a). DMACA staining of seed coat and extraction analysis further confirmed that the mutation of *NtAn1* genes blocked the specific PAs deposition in the seed coat (Fig. 3c and Fig. 3d). Most importantly, seed lipid and protein contents were both dramatically increased by targeted mutagenesis of *NtAn1* genes (Fig. 5). In *Arabidopsis*, TT8 induces the expression of late biosynthesis genes (LBGs) of PAs through directly binding to the regulatory region [29]. Natural mutation of *BrTT8* by a large insertion resulted in yellow seed coat in *Brassica rapa*, and the LBGs were significantly down-regulated by *BrTT8* mutation [22]. In this paper, the LBGs, including *ANR* and *LAR*, were significantly decreased in both *an1-1* and *an1-2* mutant lines (Fig. 3d). In *Arabidopsis*, TT8 protein could also directly bind to the promoter region of transcription factors important for seed development and lipid biosynthesis and repress the expression of them [21]. In this paper, the expression of the downstream transcription factors and fatty acid biosynthesis genes were significantly up-regulated during seed development in *an1-1* mutant line (Fig. 7). Most recently, similar phenomena were observed in CRISPR-Cas9 mediated *TT8* gene mutation in *Brassica napus* [12]. Previous reports and results in this present paper indicated that *TT8* homolog gene functions in a conserved way in seed coat PAs and seed lipid biosynthesis regulation. PAs are widely observed in seed coat in a number of oilseed plant, such as *Camelina sativa*, *Camellia oleifera*, and Tree peony. We proposed that *TT8* homolog genes could be used as an ideal target for enhancing seed lipid accumulation in these plants.

CRISPR-Cas9 system could be used for de novo tobacco domestication for biodiesel feedstock production

In recent years, the CRISPR-Cas9 based sequence-specific nucleases (SSNs) had been demonstrated to be the most simple and efficient tool for targeted gene editing. The CRISPR-Cas9 system has been successfully utilized in tobacco to generate the required mutagenesis for agronomic traits improvement and gene function characterization [30–32]. Previous reports had demonstrated that homozygous mutant could be obtained in the first generation in diverse plants, including tomato, rice, grape, and poplar [33–36]. In this present paper, two homozygous mutant lines were generated in the first generation (Fig. 2), our results demonstrated again that CRISPR-Cas9 system is a highly efficient method for targeted gene

mutation. The mutant lines showed an expected increase in seed lipid content (Fig. 5a). In addition, the seed yield was not compromised.

Due to the ever-increasing global need for energy and environmental concerns about the effects of increasing carbon dioxide levels, the demand for biofuels has been dramatically increased in the past decades [27]. To meet the huge demand for biofuel feedstocks, *de novo* domestications of plants for biofuel production have attracted lot of attentions worldwide [37, 38]. Tobacco seed oil is a promising feedstock for biodiesel production, however, current tobacco varieties are not bred for seed lipid production. With the development of CRISPR-Cas9 system, it has been proposed that CRISPR-Cas9 mediated genome editing could be used as a new tool by breeders to accelerate the domestication of semi-domesticated or even wild plants [39–41]. Low linoleic acid and high oleic acid content is a preferred character for biodiesel feedstock production, we created a high-oleic acid tobacco variety using CRISPR-Cas9 mediated *NtFAD2-2* gene editing technology in our previous work [28]. In the past few years, CRISPR-Cas9 mediated domestication had been carried out in a number of wild plant species, such as wild tomato and groundcherry [42–45].

Tobacco belongs to the Solanaceae family, which contains several well-characterized model crops, including tomato, potato, and pepper. Numerous regulators important for yield related traits, including fruit size, inflorescence and shoot architecture, had been identified and characterized in these model crop species, especially tomato [46]. These genes include *SP (SELF-PRUNING)* [47], *fw2.2 (FRUIT WEIGHT 2.2)* [48], *FASCIATED* [49], and *MULTIFLORA* [50]. Previous studies had demonstrated that the genetic regulation networks of agronomic traits were conserved in different plant taxa, which suggested that editing the homolog genes across species may generate similar phenotypes [51]. We suggested that the domestication knowledge from model crops could be translated into tobacco for the generation of high seed yield and high lipid content varieties. In this paper, seed lipid content was significantly increased by targeting a single gene, multiplex CRISPR-Cas9 system could be applied in the future to simultaneously target several genes for multiple traits enhancement.

Conclusion

In this study, we showed that *NtAn1a* and *NtAn1b* genes were highly expressed in developing tobacco seed. Targeted mutation of *NtAn1* genes were generated using the CRISPR-Cas9 mediated genome editing technology. Due to the defects in PAs biosynthesis, the mutant seeds showed the yellow-seed phenotype. We showed that targeted mutagenesis of *NtAn1* genes enhanced the seed lipid accumulation by about 18% than WT control seeds. The high knockout efficiency and significantly elevated lipid content in mutant seeds indicated that CRISPR-Cas9 could be applied to generate new tobacco varieties for biodiesel production in a faster way than traditional breeding method.

Methods

Plant material and growth condition

Wild type (WT) tobacco cultivar (*Nicotiana tabacum* L. "K326") was used for gene expression analysis and genetic transformation. Tobacco seeds were surface-sterilized with 75% ethanol for 5 min and washed three times with absolute ethanol. Then, the surface-sterilized seeds were germinated on 1/2 MS medium containing 2% sucrose and 0.7% agar at 25 ± 1 °C. The seedlings were transferred to soil and grown in greenhouse at 25 ± 1 °C with a photoperiod of 16-h light/8-h dark.

Quantitative real-time PCR (qRT-PCR)

Total RNA was isolated using Plant Total RNA Isolation Kit (Foregene, China). PrimeScript™ RT reagent Kit (Takara, Dalian) was used to generate cDNA according to the manufacturer's instruction. qRT-PCR was carried out using SYBR Premix Ex Taq™ II (Takara, Dalian) on a CFX96 real-time PCR system (Bio-Rad, USA). The qRT-PCR cycling began with one cycle at 95 °C for 2 min, followed by 40 cycles at 95 °C for 30 s, 55 °C for 30 s, 72 °C for 30 s. The house-keeping gene, *NtGAPDH* (XM_016643257), was used as the reference standard. Three biological and three technical replicates were performed. Primers used for qRT-PCR analysis were listed in Additional file 1: Table. S1.

Vector construction

To generate knockout mutants using CRISPR-Cas9 system, the pKSE401 vector (Purchased from Addgene: #62202) was used for tobacco genetic transformation. Two guide RNA (gRNA) targeting the second exon of *NtAn1* coding region were designed using CRISPR-P 2.0 (<http://crispr.hzau.edu.cn/CRISPR2/>) [52]. The scaffold containing two gRNAs was amplified by PCR using pCBC-DT1T2 vector (Purchased from Addgene: #50590) as a template. The PCR product was purified using the Universal DNA Purification Kit (TIANGEN, China) and inserted into the pKSE401 vector using golden-gate assembly method as described previously [24]. The reconstructed vector was introduced into the DH5α strain of *Escherichia coli* and confirmed using the Sanger sequencing method (Sangon Biotech, China). The verified vector (Named as *pKSE401-An1*) was introduced into *Agrobacterium tumefaciens* strain GV3101 for tobacco genetic transformation. Primers used for vector construction were listed in Additional file 1: Table. S2.

Tobacco transformation and mutant selection

Agrobacterium-mediated tobacco leaf disc transformation experiment was performed as previously described [53]. T0 generation transgenic lines were selected on MS medium supplemented with 50 mg/L kanamycin. Genomic DNA was extracted from the kanamycin-resistant T0 transgenic lines using Super Plant Genomic DNA Kit (TIANGEN, China). To select mutant lines, the flanking region of the gRNAs targeting sites was amplified using sequence specific primers by PCR (Primers were listed in Additional file 1: Table. S2). The PCR products were purified and sequenced immediately by Sanger method (Sangon Biotech, China). The T0 mutant lines were self-pollinated to generate T1 seeds. The T1 plants were analyzed again to confirm the mutation. In addition, the presence of the CRISPR-Cas9 construct in T1 mutant plants were examined by PCR using vector specific primers (Primers sequence were listed in Additional file 1: Table. S2). The homozygous T1 mutant plants without CRISPR-Cas9 construct were used for further analysis.

DMACA staining

To visualize the accumulation of PAs in seed coat, dry mature tobacco seeds from WT and mutant plants were stained in a freshly prepared dimethylaminocinnamaldehyde (DMACA, Sigma, USA) reagent [2% (w/v) DMACA dissolved in 6 N HCl/95% ethanol mixture (1:1, v/v)] for 30 min and then washed several times with 70% ethanol (v/v) as described previously [54]. The stained seeds were photographed using a Leica stereomicroscope (Leica, Germany).

PAs quantification

To extract soluble PAs from tobacco seeds, 200 mg of dry seeds were ground in liquid nitrogen and extracted with 1 ml extraction solution (70% acetone/0.5% acetic acid) by vortexing for 10 s. After sonication at room temperature for 1 h, the mixture was centrifugation at 2,500 g for 10 min, and the residue was re-extracted twice. The pooled supernatants were extracted twice with hexane. To quantify the soluble PAs level, 50 μ L of the supernatant sample was mixed with 200 μ L of DMACA reagent (0.1% DMACA, 90% ethanol, 10% HCl) in 96-well plates and the absorption was measured at 640 nm. Soluble PA levels were calculated using a standard curve prepared using procyanidin B1 (Sigma, USA).

The residue from soluble PAs extraction was air dried and used for quantitative analysis of insoluble PAs. 500 μ L butanol-HCl reagent (95% butanol: 5% concentrated HCl) was added to the residue and the mixture was sonicated at room temperature for 1 h, followed by centrifugation at 2,500 g for 10 min. The absorption of the supernatant was measured at 550 nm, then samples were boiled for 1 h, cooled to room temperature, and the absorbance at 550 nm was recorded again, with the first value being subtracted from the second. Absorbance values were converted into PAs equivalents using a standard curve of procyanidin B1 (Sigma, USA).

Seed lipid assay

To determine the seed lipid content and fatty acid composition, fatty acid methyl esters (FAMES) were prepared as previously described [55]. Twenty mature tobacco seeds were added into the methyl esterification solution [1 mL of 5% sulfuric acid in methanol (v/v), 25 μ L 0.2% butylated hydroxyl toluene solution, and 300 μ L toluene. 20 μ g triheptadecanoin was added as internal standard]. The mixture was heated at 90 °C for 2 h. Then, 1.5 mL of 0.9% NaCl and 1 mL hexane were added after the mixture had cooled down to room temperature. The FAMES were separated by collecting the organic phase. The FAMES were quantitatively analyzed by gas chromatography mass spectrometry (GC-MS) (DAOJING, Japan). The GC conditions were as follows: 1 μ L injection volume, split injection (1: 20), injector temperature 220 °C, oven temperature program: 150 °C for 1 min, then increased to 200 °C at 10 °C min⁻¹, holding at 200 °C for 1 min, then increased to 210 °C at 5 °C min⁻¹ and held for 1 min.

Protein assay

To evaluate the seed protein content, total protein were extracted as previously described with some modification [56]. Briefly, ten mature tobacco seeds were grounded in protein extraction solution [63 mM

Tris buffer, pH 7.8, 0.5 M NaCl, and 0.07% (v/v) β -mercaptoethanol]. The homogenate samples were centrifuged at 12,000 rpm for 10 min. 20 μ L of the supernatant was used for protein quantification using the Pierce BCA Protein Assay Kit (Thermo, USA) according to the manufacturer's protocol. The experiment was repeated three times.

Yield related traits assay

For seed size measurement, mature seeds were photographed using a Leica stereomicroscope (Leica, Germany). The seed length and width were measured with the Image J software. For average seed weight analysis, fifty seeds were randomly collected and carefully weighed using an electronic balance (METTLER TOLEDO, USA). Average seed weight was calculated by dividing the total seed weight by the seed number. Fruit number per plant was checked from 20 individual plants at the mature stage. Fruits used for seed number analysis were obtained from the first five basal fruits of the main inflorescence. The total seed weight from a single fruit was weighed using an electronic balance, and the seed number per fruit was calculated by dividing the total seed weight by the average seed weight.

Abbreviations

CRISPR: Clustered regularly interspaced short palindromic repeats

Cas9: CRISPR associated protein 9

DAF: Days after flowering

DMACA: Dimethylaminocinnamaldehyde

FAME: Fatty acid methyl ester

GC-MS: Gas chromatography-mass spectrometry

LBG: Late biosynthesis gene

LEC: LEAFY COTYLEDON

PA: Proanthocyanidin

PAM: Protospacer adjacent motif

T-DNA: Transfer-DNA

TT8: Transparent testa 8

WT: Wild type

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analyzed during this study are included in this published article [and its supplementary information files].

Competing interests

The authors declare that they have no competing interests.

Funding

The research was supported by the Special Project for New Transgenic Technologies and Methods of Ministry of Agriculture [No. 2016ZX08010001-010] and the Special Fund for Strategic Cooperation between Panzhihua and Sichuan University (No. 2018CDPZH-11). The funders had no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Authors' contributions

YT, and YX designed the experiments, and YT, XNL, HQ, XL, and KC performed the experiments. TL provided some guide for the experiments. All authors reviewed, discussed, and interpreted the results. YT and YX wrote the manuscript, and FC reviewed the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Usta N, Aydoğan B, Çon AH, Uğuzdoğan E, Özkal SG. Properties and quality verification of biodiesel produced from tobacco seed oil. *Energy Conversion and Management*. 2011; 52(5):2031-9.
2. Zlatanov M, Angelova M, Antova G. Lipid composition of tobacco seeds. *Bulgarian Journal of Agricultural Science*. 2007;38(1):39-40.
3. Giannelos PN, Zannikos F, Stournas S, Lois E, Anastopoulos G. Tobacco seed oil as an alternative diesel fuel: physical and chemical properties. *Industrial Crops & Products*. 2002; 16(1):1-9.
4. Veljkovic V, Lakicevic S, Stamenkovic O, Todorovic Z, Lazic M. Biodiesel production from tobacco (*Nicotiana tabacum* L.) seed oil with a high content of free fatty acids. *Fuel*. 2006; 85(17-18):2671-5.
5. kumar M, Sharma MP. Selection of potential oils for biodiesel production. *Renewable and Sustainable Energy Reviews*. 2016; 56:1129-38.
6. Banković-Ilić IB, Stamenković OS, Veljković VB. Biodiesel production from non-edible plant oils. *Renewable and Sustainable Energy Reviews*. 2012; 16(6):3621-47.
7. Zhao, Min, Zhang, Ke, Wang, Jing, Chen, Chaoying, Cai, Ji Xiaoxiao. Hydrothermal carbonization of tobacco stalk for fuel application. *Bioresource Technology*. 2016; 220, 305-11. .
8. Grisan S, Polizzotto R, Raiola P, Cristiani S, Ventura F, di Lucia F, Zuin M, Tommasini S, Morbidelli R, Damiani F et al. Alternative use of tobacco as a sustainable crop for seed oil, biofuel, and biomass. *Agronomy for Sustainable Development*. 2016; 36(4):55.
9. Carvalho FS, Fornasier F, Leitão JOM, Moraes JAR, Schneider RCS. Life cycle assessment of biodiesel production from solaris seed tobacco. *Journal of Cleaner Production*. 2019; 230:1085-95.
10. Poltronieri P. Tobacco Seed Oil for Biofuels. 2016;161-87.
11. Wang Z, Chen M, Chen T, Xuan L, Li Z, Du X, Zhou L, Zhang G, Jiang L. TRANSPARENT TESTA2 regulates embryonic fatty acid biosynthesis by targeting *FUSCA3* during the early developmental stage of *Arabidopsis* seeds. *Plant J*. 2014; 77(5):757-69.
12. Zhai Y, Yu K, Cai S, Hu L, Amoo O, Xu L, Yang Y, Ma B, Jiao Y, Zhang C et al. Targeted mutagenesis of *BnTT8* homologs controls yellow seed coat development for effective oil production in *Brassica napus* L. *Plant Biotechnol J*. 2019; 18: 1153-68.
13. Lepiniec L, Debeaujon I, Routaboul JM, Baudry A, Pourcel L, Nesi N, Caboche M. Genetics and biochemistry of seed flavonoids. *Annual review of plant biology*. 2006; 57:405-30.
14. Xu W, Dubos C, Lepiniec L. Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. *Trends Plant Sci*. 2015; 20(3):176-85.
15. Baudry A, Heim MA, Dubreucq B, Caboche M, Weisshaar B, Lepiniec L. TT2, TT8, and TTG1 synergistically specify the expression of *BANYULS* and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *Plant J*. 2004; 39(3):366-80.
16. Nesi N, Debeaujon I, Jond C, Pelletier G, Caboche M, Lepiniec L. The *TT8* gene encodes a basic helix-loop-helix domain protein required for expression of *DFR* and *BAN* genes in *Arabidopsis* siliques. *Plant Cell*. 2000; 12(10):1863-78.

17. Padmaja LK, Agarwal P, Gupta V, Mukhopadhyay A, Sodhi YS, Pental D, Pradhan AK. Natural mutations in two homoeologous *TT8* genes control yellow seed coat trait in allotetraploid *Brassica juncea* (AABB). *Theor Appl Genet.* 2014; 127(2):339-47.
18. Li P, Chen B, Zhang G, Chen L, Dong Q, Wen J, Mysore KS, Zhao J. Regulation of anthocyanin and proanthocyanidin biosynthesis by *Medicago truncatula* bHLH transcription factor MtTT8. *New Phytol.* 2016; 210(3):905-21.
19. Escaray FJ, Passeri V, Perea-Garcia A, Antonelli CJ, Damiani F, Ruiz OA, Paolocci F. The R2R3-MYB *TT2b* and the bHLH *TT8* genes are the major regulators of proanthocyanidin biosynthesis in the leaves of *Lotus* species. *Planta.* 2017; 246(2):243-61.
20. Lim SH, Kim DH, Kim JK, Lee JY, Ha SH. A radish basic helix-loop-helix transcription factor, RsTT8 acts a positive regulator for anthocyanin biosynthesis. *Front Plant Sci.* 2017; 8:1917.
21. Chen M, Xuan L, Wang Z, Zhou L, Li Z, Du X, Ali E, Zhang G, Jiang L. TRANSPARENT TESTA8 inhibits seed fatty acid accumulation by targeting several seed development regulators in *Arabidopsis*. *Plant Physiol.* 2014; 165(2):905-16.
22. Li X, Chen L, Hong M, Zhang Y, Zu F, Wen J, Yi B, Ma C, Shen J, Tu J et al. A large insertion in bHLH transcription factor *BrTT8* resulting in yellow seed coat in *Brassica rapa*. *PLoS One.* 2012; 7(9):e44145.
23. Bai Y, Pattanaik S, Patra B, Werkman JR, Xie CH, Yuan L. Flavonoid-related basic helix-loop-helix regulators, *NtAn1a* and *NtAn1b*, of tobacco have originated from two ancestors and are functionally active. *Planta.* 2011; 234(2):363-75.
24. Xing HL, Dong L, Wang Z-P, Zhang H-Y, Han C-Y, Liu B, Wang X-C, Chen Q-J. A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *Bmc Plant Biology.* 2014; 14(1):327.
25. Xie DY, Sharma SB, Paiva NL, Ferreira D, Dixon RA. Role of anthocyanidin reductase, encoded by *BANYULS* in plant flavonoid biosynthesis. *Science.* 2003; 299(5605):396-9.
26. Tanner GJ, Francki KT, Abrahams S, Watson JM, Larkin PJ, Ashton AR. Proanthocyanidin biosynthesis in plants. *Journal of Biological Chemistry.* 2003; 278(34):31647-56.
27. Durrett TP, Benning C, Ohlrogge J. Plant triacylglycerols as feedstocks for the production of biofuels. *Plant J.* 2008; 54(4):593-607.
28. Tian Y, Chen K, Li X, Zheng Y, Chen F. Design of high-oleic tobacco (*Nicotiana tabacum* L.) seed oil by CRISPR-Cas9-mediated knockout of *NtFAD2-2*. *BMC Plant Biology.* 2020; 20(1):233.
29. Xu W, Grain D, Bobet S, Gourrierc JL, Thévenin J, Kelemen Z, Lepiniec L, Dubos C. Complexity and robustness of the flavonoid transcriptional regulatory network revealed by comprehensive analyses of MYB–bHLH–WDR complexes and their targets in *Arabidopsis* seed. *New Phytol.* 2014; 202(1):132-44.
30. Schachtsiek J, Stehle F: Nicotine-free, nontransgenic tobacco (*Nicotiana tabacum* L.) edited by CRISPR-Cas9. *Plant Biotechnology Journal* 2019; 17(12): 2228-30.
31. Gao J, Tong Z, Xu B, Ling J, Xiao B, He L, Liu L, Hao Y, Xia Q. CRISPR/Cas9-Mediated mutagenesis of *carotenoid cleavage dioxygenase 8 (CCD8)* in tobacco affects shoot and root architecture.

- International Journal of Molecular Sciences. 2018; 19(4):1062.
32. Jiang W, Zhou H, Bi H, Michael F, Bing Y, Weeks DP. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Research*. 2013; 41(20): e188.
 33. Pan C, Ye L, Qin L, Liu X, He Y, Wang J, Chen L, Lu G. CRISPR/Cas9-mediated efficient and heritable targeted mutagenesis in tomato plants in the first and later generations. *Scientific reports*. 2016; 6:24765.
 34. Fan D, Liu T, Li C, Jiao B, Li S, Hou Y, Luo K. Efficient CRISPR/Cas9-mediated Targeted Mutagenesis in Populus in the First Generation. *Scientific reports*. 2015; 5:12217.
 35. Zhang H, Zhang J, Wei P, Zhang B, Gou F, Feng Z, Mao Y, Yang L, Zhang H, Xu N et al. The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant Biotechnol J*. 2014; 12(6):797-807.
 36. Wang X, Tu M, Wang D, Liu J, Li Y, Li Z, Wang Y, Wang X: CRISPR/Cas9-mediated efficient targeted mutagenesis in grape in the first generation. *Plant Biotechnology Journal*. 2018; 16(4):844-55.
 37. Montes JM, Melchinger AE. Domestication and breeding of *Jatropha curcas* L. *Trends in plant science*. 2016; 21(12):1045-57.
 38. Sainger M, Jaiwal A, Sainger PA, Chaudhary D, Jaiwal R, Jaiwal PK. Advances in genetic improvement of *Camelina sativa* for biofuel and industrial bio-products. *Renewable and Sustainable Energy Reviews*. 2017; 68:623-37.
 39. Osterberg JT, Xiang W, Olsen LI, Edenbrandt AK, Vedel SE, Christiansen A, Landes X, Andersen MM, Pagh P, Sandoe P et al. Accelerating the domestication of new crops: feasibility and approaches. *Trends Plant Sci*. 2017; 22(5):373-84.
 40. Fernie AR, Yan J. *De novo* domestication: an alternative route toward new crops for the future. *Mol Plant*. 2019; 12(5):615-31.
 41. Khan MZ, Zaidi SS, Amin I, Mansoor S. A CRISPR way for fast-forward crop domestication. *Trends Plant Sci*. 2019; 24(4):293-6.
 42. Lemmon ZH, Reem NT, Dalrymple J, Soyk S, Swartwood KE, Rodriguez-Leal D, Van Eck J, Lippman ZB. Rapid improvement of domestication traits in an orphan crop by genome editing. *Nat Plants*. 2018; 4(10):766-70.
 43. Zsogon A, Cermak T, Naves ER, Notini MM, Edel KH, Weini S, Freschi L, Voytas DF, Kudla J, Peres LEP. *De novo* domestication of wild tomato using genome editing. *Nat Biotechnol*. 2018; 36, 1211-6.
 44. Li T, Yang X, Yu Y, Si X, Zhai X, Zhang H, Dong W, Gao C, Xu C. Domestication of wild tomato is accelerated by genome editing. *Nat Biotechnol*. 2018; 36, 1160-3.
 45. Zsogon A, Cermak T, Voytas D, Peres LE. Genome editing as a tool to achieve the crop ideotype and *de novo* domestication of wild relatives: Case study in tomato. *Plant Sci*. 2017; 256:120-30.
 46. Rothan C, Diouf I, Causse M. Trait discovery and editing in tomato. *Plant J*. 2019; 97(1):73-90.

47. Pnueli L, Carmel-Goren L, Hareven D, Gutfinger T, Alvarez J, Ganai M, Zamir D, Lifschitz E. The *SELF-PRUNING* gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of *CEN* and *TFL1*. *Development*. 1998; 125(11):1979-89.
48. Frary A, Nesbitt TC, Grandillo S, Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB, Tanksley SD. *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. *Science*. 2000; 289(5476):85-8.
49. Xu C, Liberatore KL, MacAlister CA, Huang Z, Chu YH, Jiang K, Brooks C, Ogawa-Ohnishi M, Xiong G, Pauly M et al. A cascade of arabinosyltransferases controls shoot meristem size in tomato. *Nat Genet*. 2015; 47(7):784-92.
50. Lippman ZB, Cohen O, Alvarez JP, Abu-Abied M, Pekker I, Paran I, Eshed Y, Zamir D. The making of a compound inflorescence in tomato and related nightshades. *PLoS Biol*. 2008; 6(11):e288.
51. Hussain Q, Shi J, Scheben A, Zhan J, Wang X, Liu G, Yan G, King GJ, Edwards D, Wang H. Genetic and signalling pathways of dry fruit size: targets for genome editing-based crop improvement. *Plant Biotechnol J*. 2020; 18(5):1124-40.
52. Liu H, Ding Y, Zhou Y, Jin W, Xie K, Chen LL. CRISPR-P 2.0: an improved CRISPR-Cas9 tool for genome editing in plants. *Molecular Plant*. 2017; 10(3):530-2.
53. Horsch RB. A simple and general method for transferring genes into plants. *Science*. 1985; 227(4691):1229-31.
54. Hong M, Hu K, Tian T, Li X, Chen L, Zhang Y, Yi B, Wen J, Ma C, Shen J et al. Transcriptomic analysis of seed coats in yellow-seeded *Brassica napus* reveals novel genes that influence proanthocyanidin biosynthesis. *Front Plant Sci*. 2017; 8:1674.
55. Li Y, Beisson F, Pollard M, Ohlrogge J. Oil content of *Arabidopsis* seeds: the influence of seed anatomy, light and plant-to-plant variation. *Phytochemistry*. 2006; 67(9):904-15.
56. Heath JD, Weldon R, Monnot C, Meinke DW. Analysis of storage proteins in normal and aborted seeds from embryo-lethal mutants of *Arabidopsis thaliana*. *Planta*. 1986; 169(3):304-12.

Figures

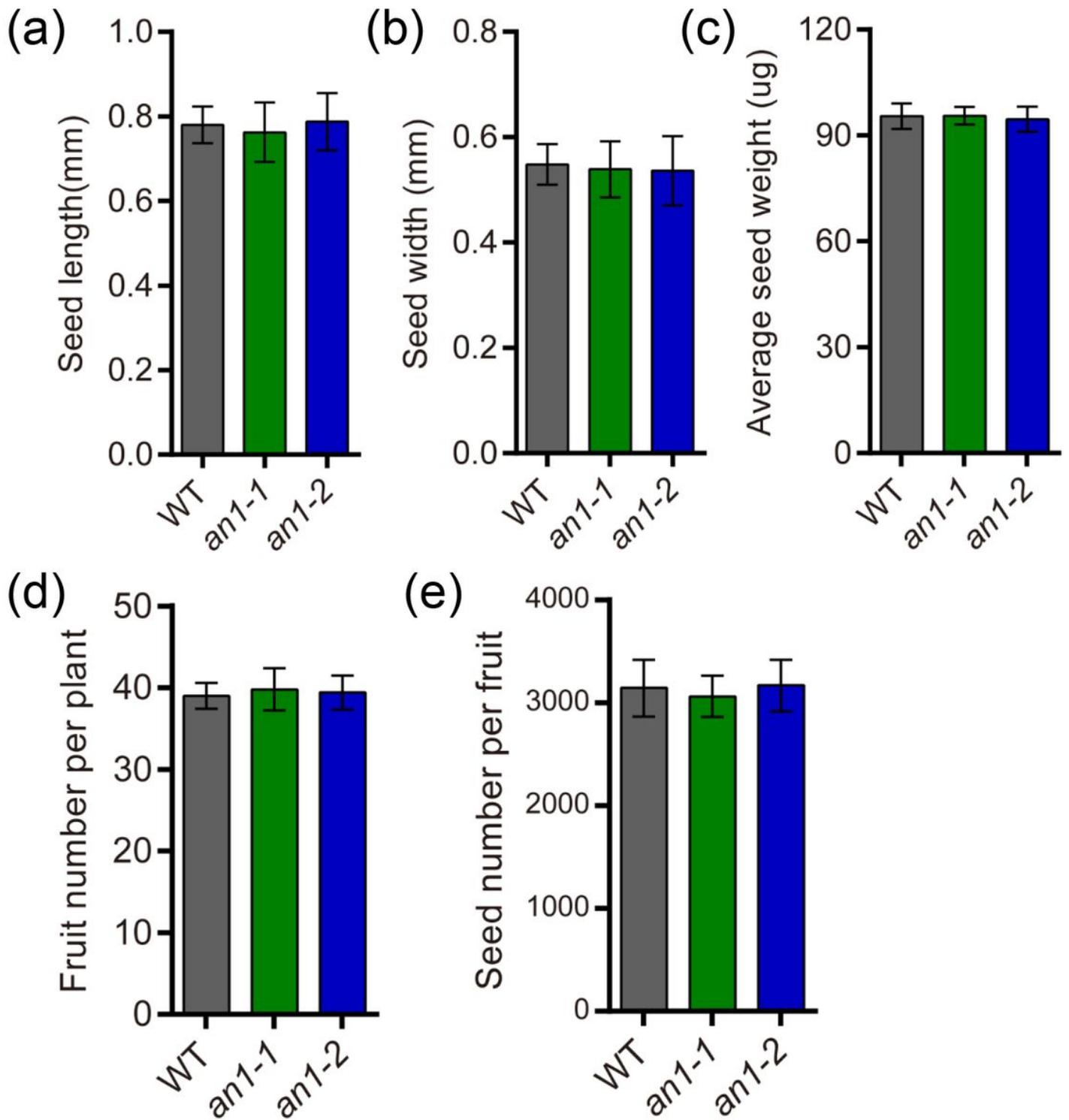


Figure 1

Seed yield related traits of an1 mutant tobacco lines. (a) Average seed length. Data are mean \pm SD (n = 30). (b) Average seed width. Data are mean \pm SD (n = 30). (c) Average seed weight. Values are mean \pm SD of five individual measurements of 50 seeds/replicate. The seeds used to measure seed size and seed weight were selected randomly after collecting all of the mature seeds. (d) Fruit number per plant. Fruit number was calculated at the plant mature stage. Data are mean \pm SD (n = 20). (e) Seed number per fruit.

Data are mean \pm SD (n = 10). Seed number per fruit was calculated from the basal five fruits at the mature stage.

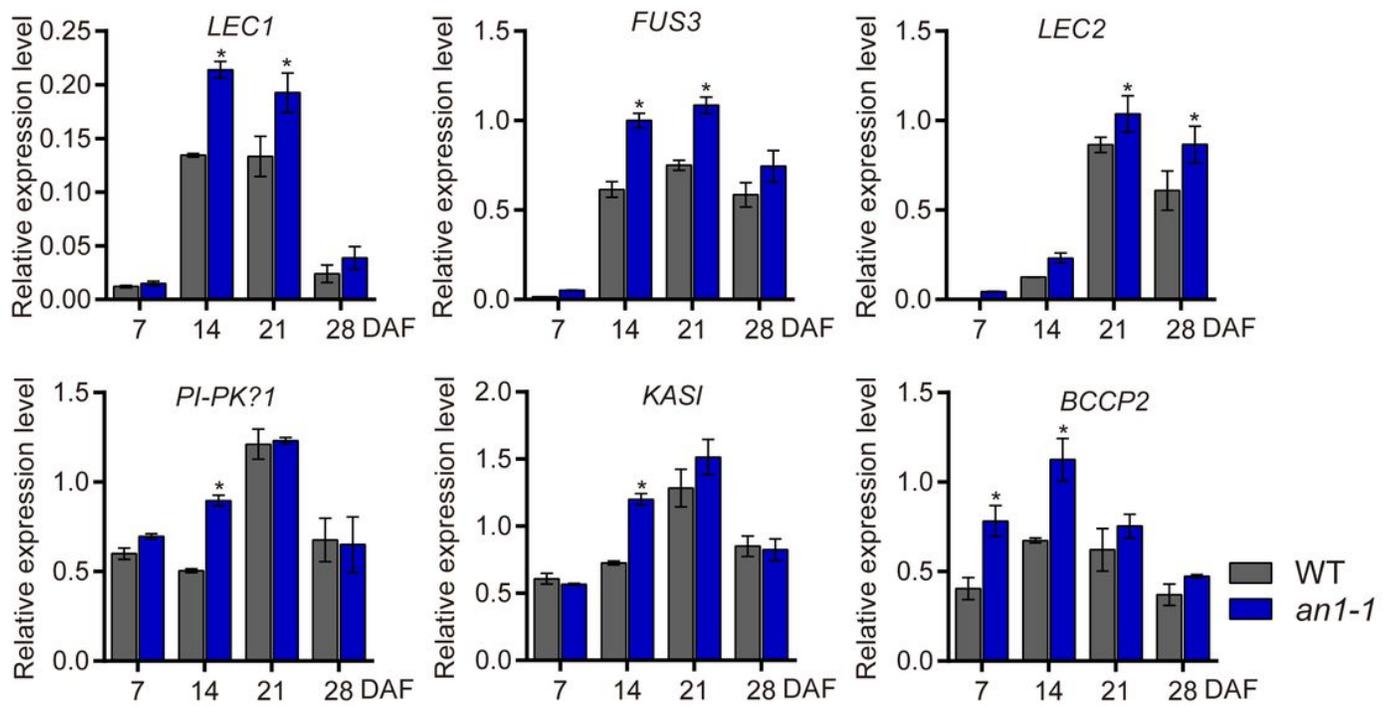


Figure 2

Expression analysis of genes involved in seed development and fatty acid biosynthesis. Asterisks indicate significant difference compared to control samples (*P < 0.05).

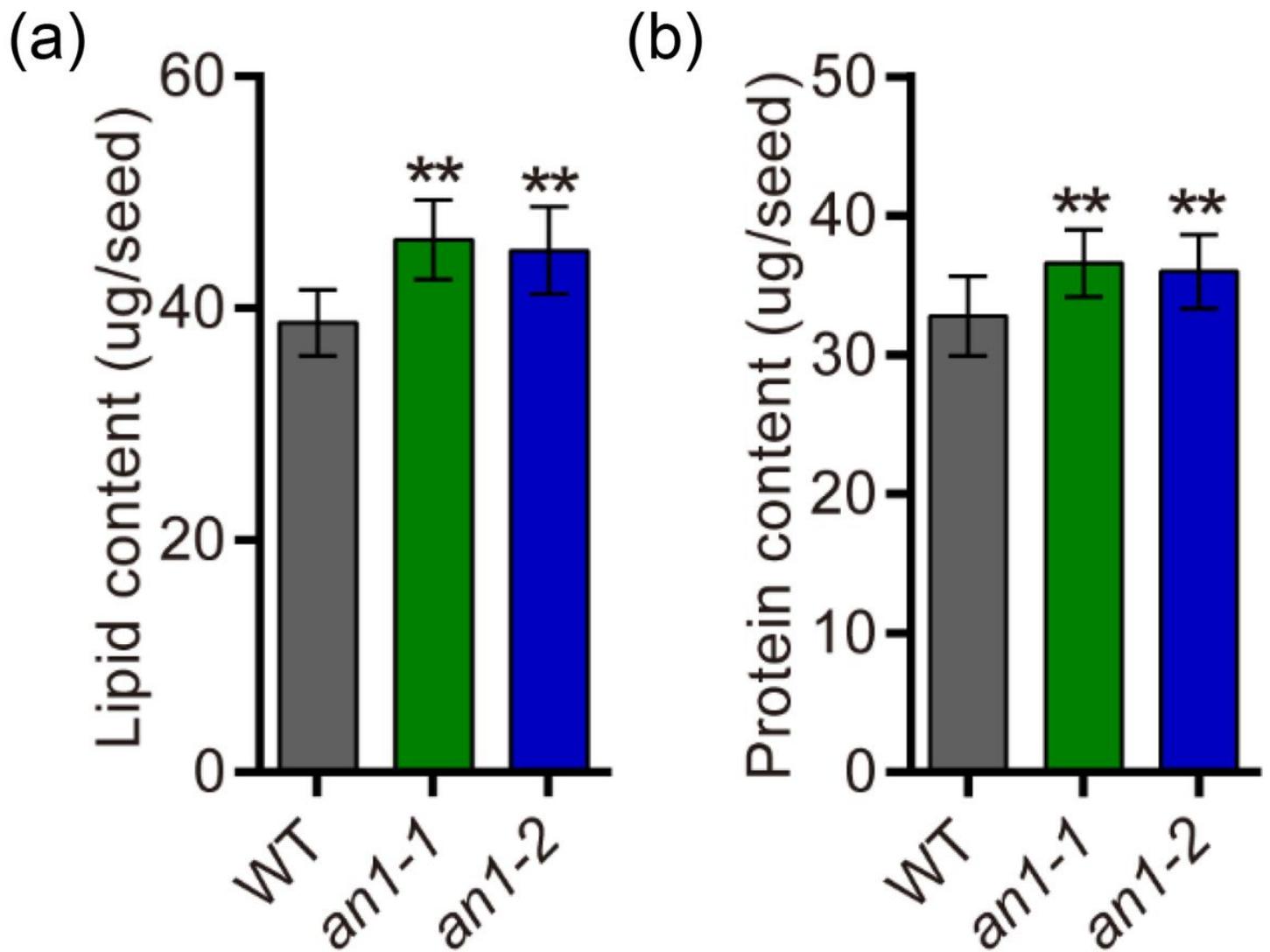


Figure 3

Lipid and protein content of the WT and an1 mutant seeds. (a) Lipid content per seed. Twenty seeds were analyzed for each repeat and per seed lipid content was calculated by dividing the seed number. Data are mean \pm SD (n = 3). (b) Protein content per seed. Ten seeds were analyzed for each repeat and per seed protein content was calculated by dividing the seed number. Data are mean \pm SD (n = 3). Asterisks indicate significant difference compared to control samples. (**P < 0.01).

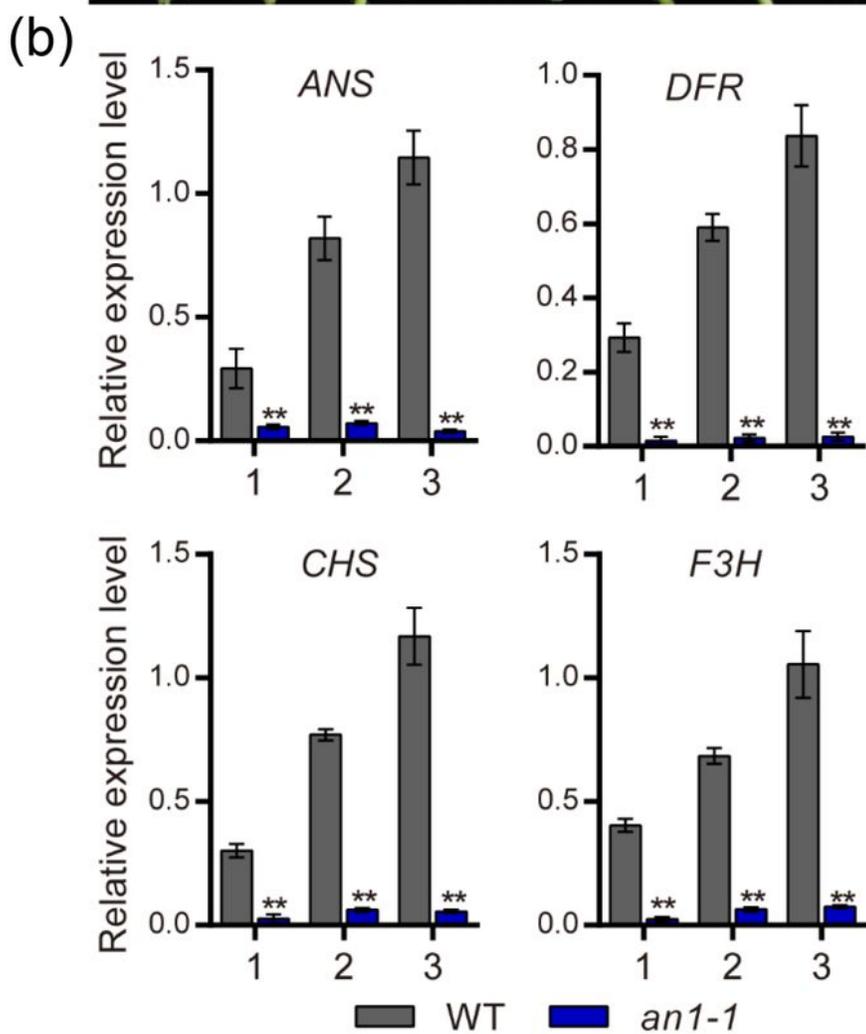


Figure 4

Flower phenotype of tobacco *an1-1* mutant. (a) Flower at different development stages from WT and *an1-1* mutant line. (b) Expression analysis of anthocyanin biosynthetic genes regulated by NtAn1 in tobacco. CHS, chalcone synthase; F3H, flavanone 3-hydroxylase; DFR, dihydroflavonol 4-reductase; ANS, anthocyanidin synthase. Data are mean \pm SD (n = 3). Asterisks indicate significant difference compared to control samples (*P < 0.05; **P < 0.01). NtGAPDH gene was used as an internal reference.

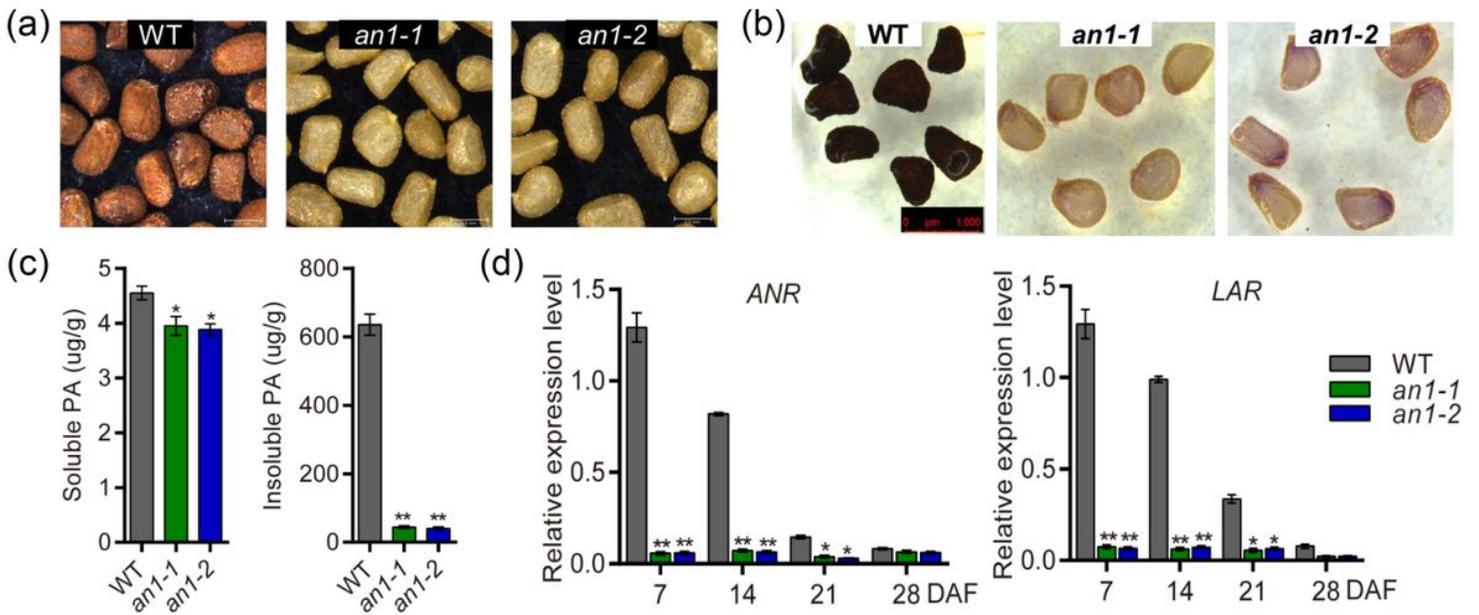


Figure 5

Seed phenotypes of tobacco *an1* mutants. (a) Tobacco seed coat color phenotypes of WT and *an1* mutant lines. (b) DMACA staining of mature tobacco seeds from WT and *an1* mutant lines. (c) Soluble and insoluble PAs contents in mature seeds of WT control and *an1* mutant lines. Three biological replicates were analyzed. Data are mean \pm SD ($n = 3$). (d) Expression of PAs biosynthetic genes regulated by NtAn1 in tobacco. ANR: anthocyanidin reductase, LAR: leucoanthocyanidin reductase. RNA samples were extracted from the developing seeds at 7, 14, 21, and 28 DAF. Data are mean \pm SD of three biological replicates. Asterisks indicate significant difference compared to control samples (* $P < 0.05$; ** $P < 0.01$). NtGAPDH gene was used as an internal reference.

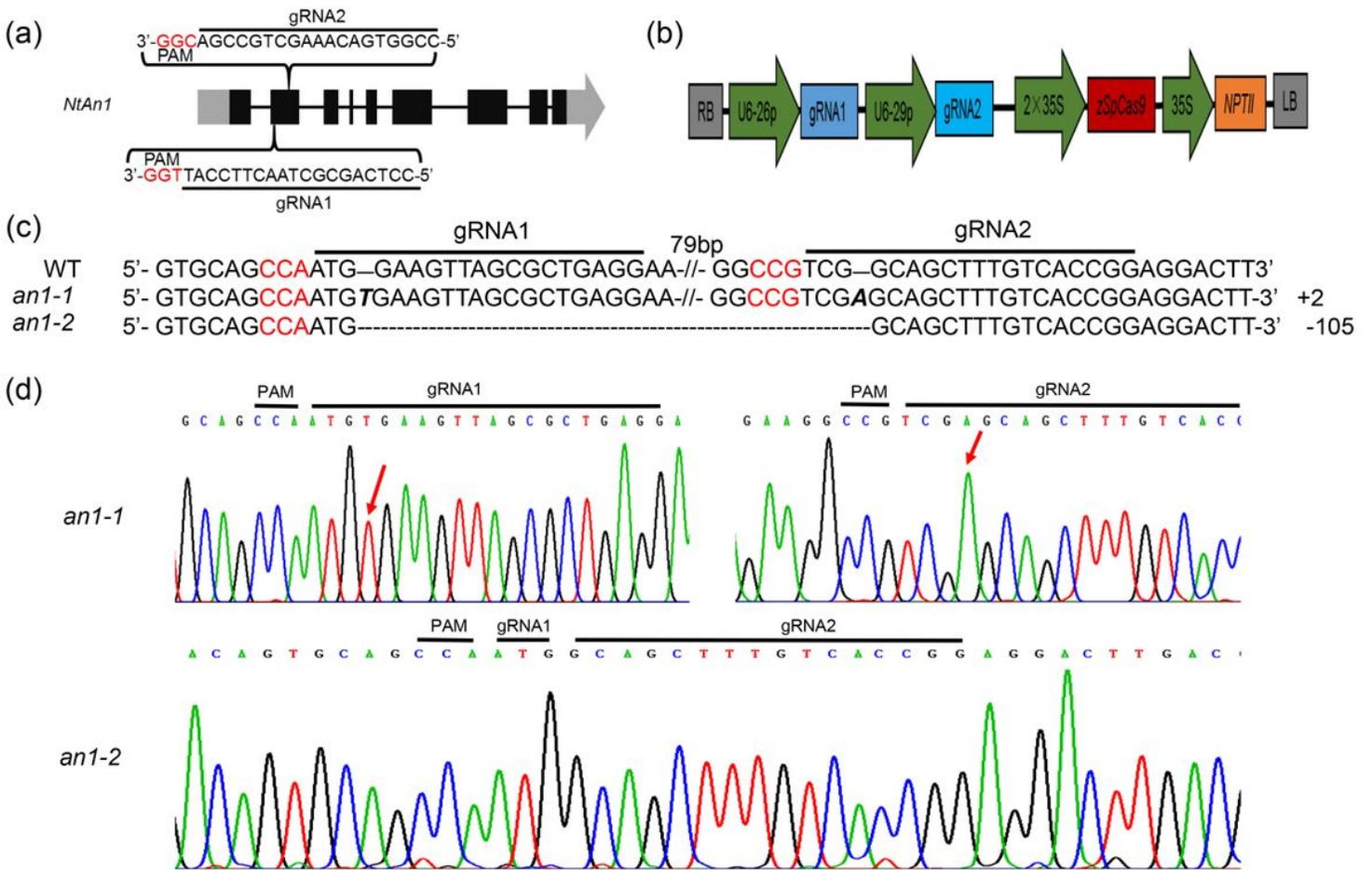


Figure 6

CRISPR-Cas9 mediated targeted mutation of *NtAn1*. (a) Targeting sites and gRNA sequences for *NtAn1* genes editing. The 3 base protospacer adjacent motif (PAM) was in red. (b) The CRISPR-Cas9 vector structure used for targeted mutation of *NtAn1*. (c) Mutation type of the *an1-1* and *an1-2* lines determined by Sanger method. (d) Sanger sequencing peaks of the *an1-1* and *an1-2* mutant lines. The red arrows indicate the location of mutations.

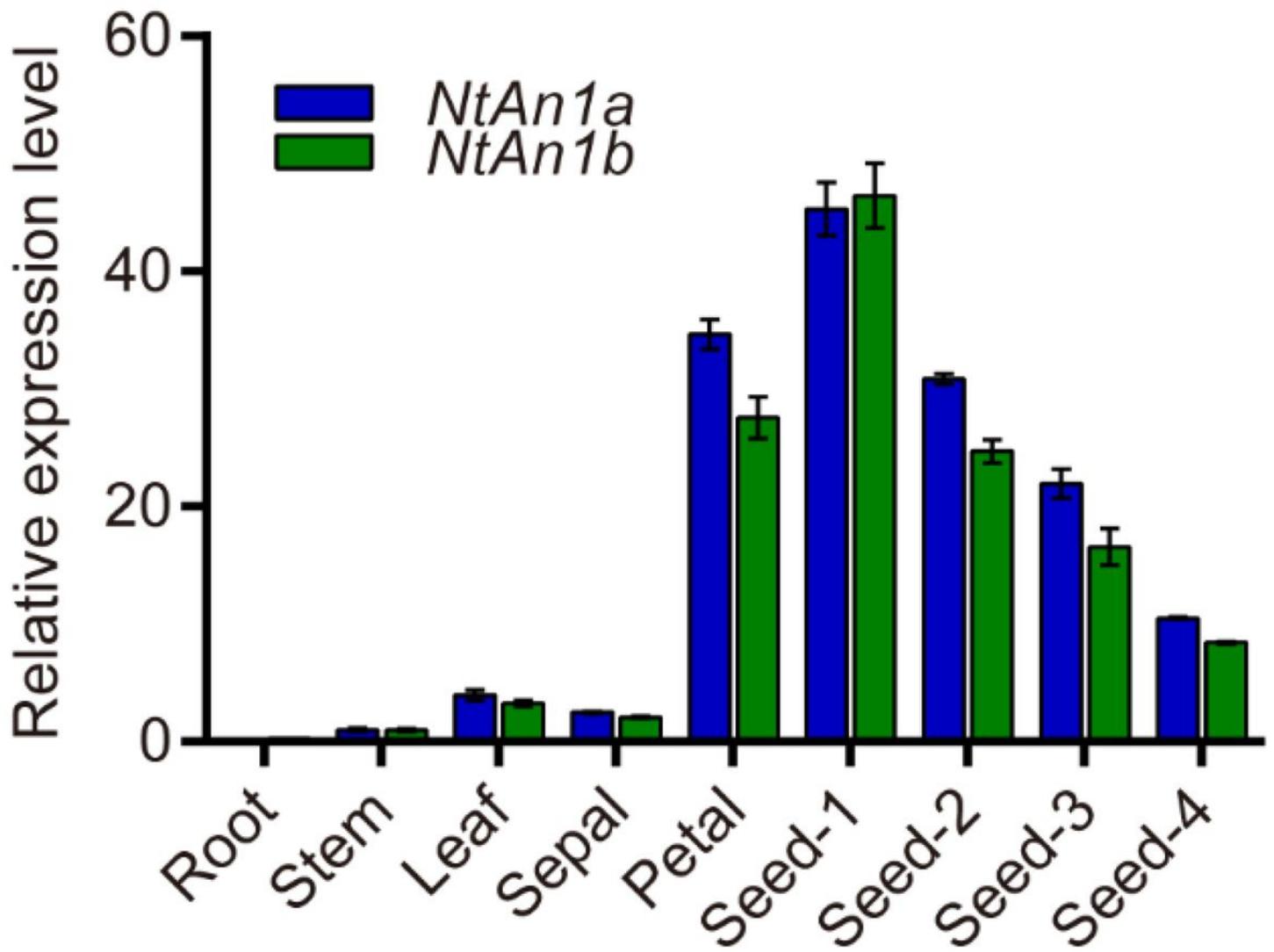


Figure 7

Expression profiles of tobacco *NtAn1a* and *NtAn1b* genes. Seed-1, developing seed at 7 days after flowering (DAF); Seed-2, developing seed at 15 DAF; Seed-3, developing seed at 21 DAF; Seed-4, developing seed at 21 DAF. Data are mean \pm SD (n = 3). *NtGAPDH* gene was amplified as an internal control.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1.doc](#)